Sign in

Web Images Video New! News Maps more »

geocoded database <1998 Search Advanced Search Preferences

Web Results **21** - **30** of about **62,600** for **geocoded database <1998**. (**0.09** seconds)

Position based personal digital assistant - Patent 6266612
A computer assisted method of using a **geocoded database**, comprising the steps of: ...
Database 10 is preferably a **geocoded database** and will be described in ...
www.freepatentsonline.com/6266612.html - 60k - Cached - Similar pages

Truck Traffic Changes in Grain Transportation
In the **1998**-99 Crop Year, the 186 super elevator locations moved 44.8% of the ... in the form illustrated in Figure 20, are provided as a **geocoded database**. ...
www.tc.gc.ca/pol/en/Report/truckGrain1999/C4.htm - 75k - Cached - Similar pages

[PDF] White Paper
File Format: PDF/Adobe Acrobat - View as HTML
geocoded. A **database** is **geocoded** first to block-face* ... In September **1998** E0C, E0H. and E0J were converted, in March 1999 E0K and ...
www.dmtispatial.com/pdfs/street_intersection.pdf - Similar pages

MassGIS - MGIC Meeting - September 1, **1998**
geocoded breast cancer cases (approx 2200); groundwater contours/census tracts ... valuable accomplishment thus far is the development of the GIS **database**. ...
www.mass.gov/mgis/mgic0998.htm - 23k - Cached - Similar pages

[PDF] Many Benefits Realized through Cadastral GIS
File Format: PDF/Adobe Acrobat
at CATIE, the City obtained and **geocoded** external ... By the end of **1998**, CATIE had organized. the **database** system and linked graphic and literal ...
www.esri.com/library/reprints/pdfs/arcuser_cadastralgis.pdf - Similar pages

Memo To Members: 3.33 9/11/**1998**
HUD will be sponsoring three public forums on October 1, **1998**, to receive comments on the ... The **database** has been **geocoded** as well. Go to www.huduser.org, ...
www.nlihc.org/mtm/mtm3-33.htm - 20k - Cached - Similar pages

[PDF] Mapping of permafrost on Vega Island, Antarctic Peninsula, using ...
File Format: PDF/Adobe Acrobat
of the northern Larsen Ice Shelf (Rott andothers,**1998**) anda. 1.5 C rise in air temperature during ... a **geocoded database**. Difficulties were found during ...
www.ingentaconnect.com/content/igsoc/agl/2002/00000034/00000001/art00028 - Similar pages

Existing Satellite Imagery **Database** for NPS
MRLC **Geocoded** Data: http://edc.usgs.gov/webglis Data from approximately 1988 to around 1993 or **1998** available. CORONA: http://edc.usgs.gov/webglis ...
www.nps.gov/gis/nps_imagery_data_base.htm - 13k - Cached - Similar pages

Location based Services (LBS) - Company Directory
Description: Established in March, **1998**, Pharos Science and Applications, ... Our showcase product is a first in the US - a **geocoded database** ... ...
www.directionsmag.com/companies/category/Location_based_Services_(LBS)/ - 115k - Cached - Similar pages

Use of Computerized Crime Mapping by Law Enforcement in the United ...
Surveys were accepted until May 1, **1998**. It took responding agencies an ... types of data **geocoded** and mapped, types of spatial analyses performed and how ...
webapp.icpsr.umich.edu/cocoon/ICPSR-STUDY/02878.xml - 27k - Cached - Similar pages

◀ Goooooooooooogle ▶
Result Page: **Previous** 1 2 3 4 5 6 7 8 9 10 11 12 **Next**

Free! Speed up the web. Download the Google Web Accelerator.

---

geocoded database <1998 [Search]

Search within results | Language Tools | Search Tips

Google Home - Advertising Programs - Business Solutions - About Google

PORTAL
USPTO

Subscribe (Full Service) Register (Limited Service, Free) Login

Search: ◉ The ACM Digital Library ○ The Guide

+geocoding centroid

SEARCH

THE ACM DIGITAL LIBRARY

Feedback Report a problem Satisfaction survey

Terms used **geocoding centroid**

Found **39** of **185,942**

Sort results by: relevance

Display results: expanded form

Save results to a Binder

Search Tips

☐ Open results in a new window

Try an Advanced Search

Try this search in The ACM Guide

Results 1 - 20 of 39

Result page: **1** 2 next

Relevance scale

1 Geographic information systems education for non-computer oriented college students
Harold G. Campbell
September 1994 **ACM SIGCSE Bulletin**, Volume 26 Issue 3
Publisher: ACM Press
Full text available: pdf(443.52 KB) Additional Information: full citation, index terms

2 Mobility support and location awareness: Developing spatially-aware content management systems for dynamic, location-specific information in mobile environments
Harsha Tummala, Joel Jones
September 2005 **Proceedings of the 3rd ACM international workshop on Wireless mobile applications and services on WLAN hotspots WMASH '05**
Publisher: ACM Press
Full text available: pdf(1.06 MB) Additional Information: full citation, abstract, references, index terms

Current location-aware information systems lack an effective method of maintaining and updating dynamic, location-specific content. We have developed a design for representing location-specific content that balances flexibility and comprehensibility. We have developed a web-based content management system that implements this design. The system provides an easy-to-use interface to tie any form of media-such as text, pictures, audio, or video-to a location.This work is directly applicable to vari ...

Keywords: content management, context-aware services, location-aware applications, mobile computing, user-driven information systems

3 Bootstrapping toponym classifiers
David A. Smith, Gideon S. Mann
May 2003 **Proceedings of the HLT-NAACL 2003 workshop on Analysis of geographic references - Volume 1**
Publisher: Association for Computational Linguistics
Full text available: pdf(62.70 KB) Additional Information: full citation, abstract, references

We present minimally supervised methods for training and testing geographic name disambiguation (GND) systems. We train data-driven place name classifiers using

toponyms already disambiguated in the training text --- by such existing cues as "Nashville, Tenn." or "Springfield, MA" --- and test the system on texts where these cues have been stripped out and on hand-tagged historical texts. We experiment on three English-language corpora of varying provenance and complexity: newsfeed from the 1990 ...

4 The MAPEDIT system for automatic map digitization
H. H. Holmes, D. M. Austin, W. H. Benson
July 1974 **Proceedings of the 1st annual conference on Computer graphics and interactive techniques**
Publisher: ACM Press
Full text available: pdf(11.83 KB) Additional Information: full citation, abstract

A system for the automatic digitization of polygon boundaries is described. Digitized map files are created from a driver tape containing identification codes and approximate centroids of polygonal boundaries (e.g., census tracts), and a film image of the map. The digitizer scans on the film plane in an automatic line-following mode, producing the first stage of the map file for the editing system. The MAPEDIT system, which can be used either interactively or in batch mode, reads maps in several ...

5 Abstracts from the conference on computer graphics and interactive techniqes
September 1974 **ACM SIGGRAPH Computer Graphics**, Volume 8 Issue 3
Publisher: ACM Press
Full text available: pdf(1.10 MB) Additional Information: full citation

6 The buffalo crime mapping system: a design strategy for the display and analysis of spatially referenced crime data
Kurt E. Brassel, Jack J. Utano, Perry O. Hanson
July 1977 **ACM SIGGRAPH Computer Graphics , Proceedings of the 4th annual conference on Computer graphics and interactive techniques SIGGRAPH '77**, Volume 11 Issue 2
Publisher: ACM Press
Full text available: pdf(181.17 KB) Additional Information: full citation, abstract, references, citings

This paper presents the design strategy for a crime information system with vast capabilities for production of and experimentationwith maps. Detailed disaggregated crime information from the city of Buffalo, New York forms the primary data base for the system. This series of crime files, together with socio-economic data from the Census, are integrated into a geographic information system denoted as the Crime Analysis and Research Package (CARP). The Buffalo Crime Mapping System as a part of CA ...

Keywords: cartographic data structures, computer cartography, criminal data processing, geographic information system, information retrieval

7 Construction of the planar partition postal code map based on cadastral reGIStration
Friso Penninga, Edward Verbree, Wilko Quak, Peter van Oosterom
November 2003 **Proceedings of the 11th ACM international symposium on Advances in geographic information systems**
Publisher: ACM Press
Full text available: pdf(847.15 KB) Additional Information: full citation, abstract, references, index terms

Accurate postal code maps could play an important role within GIS as the postal code has the potential to link the address description of buildings and their location in a certain global reference system. This relationship is possible in both directions: address matching

and geocoding. These operators demand a certain mechanism in translating an exact geometric position (i.e. WGS84 coordinate) into a location indication (town, street, house number) and vice versa. As most built-up parcels are pr ...

Keywords: GIS, skeletonization, triangulation

8 Geospatial mapping and navigation of the web
Kevin S. McCurley
April 2001 **Proceedings of the 10th international conference on World Wide Web**
Publisher: ACM Press
Full text available: pdf(1.06 MB) Additional Information: full citation, references, citings, index terms

Keywords: browsers, geographic information systems, geospatial information retrieval, navigation

9 Geographic Data Processing
George Nagy, Sharad Wagle
June 1979 **ACM Computing Surveys (CSUR)**, Volume 11 Issue 2
Publisher: ACM Press
Full text available: pdf(4.20 MB) Additional Information: full citation, references, citings, index terms

10 Statistical methods II: Shrinkage estimator generalizations of Proximal Support Vector Machines
Deepak K. Agarwal
July 2002 **Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining**
Publisher: ACM Press
Full text available: pdf(1.14 MB) Additional Information: full citation, abstract, references, citings, index terms

We give a statistical interpretation of Proximal Support Vector Machines (PSVM) proposed at KDD2001 as linear approximaters to (nonlinear) Support Vector Machines (SVM). We prove that PSVM using a linear kernel is identical to ridge regression, a biased-regression method known in the statistical community for more than thirty years. Techniques from the statistical literature to estimate the tuning constant that appears in the SVM and PSVM framework are discussed. Better shrinkage strategies that ...

Keywords: Bayesian models, bias-variance tradeoff, classification, correlation, kernel, regression

11 Spatio temporal and GIS: Scalable data collection and retrieval infrastructure for digital government applications
Hanan Samet, Leana Golubchik
May 2006 **Proceedings of the 2006 international conference on Digital government research dg.o '06**
Publisher: ACM Press
Full text available: pdf(87.04 KB) Additional Information: full citation, abstract

In this paper we describe highlights of the project titled "Scalable data collection infrastructure for digital government applications" under the auspices of the Digital

Government Research Program of the National Science Foundation. Our research is focused on taking advantage of the distributed nature of data and the interaction with it. Our efforts have been directed at both the systems/theoretical and applications levels. On the systems and theoretical levels, we have continued our developme ...

12 Experience papers: experiences with open source and legacy systems: Experiences with place lab: an open source toolkit for location-aware computing
Timothy Sohn, William G. Griswold, James Scott, Anthony LaMarca, Yatin Chawathe, Ian Smith, Mike Chen
May 2006 **Proceeding of the 28th international conference on Software engineering ICSE '06**
Publisher: ACM Press
Full text available: pdf(729.93 KB) Additional Information: full citation, abstract, references, index terms

Location-based computing (LBC) is becoming increasing important in both industry and academia. A key challenge is the pervasive deployment of LBC technologies; to be effective they must run on a wide variety of client platforms, including laptops, PDAs, and mobile phones, so that location data can be acquired anywhere and accessed by any application. Moreover, as a nascent area, LBC is experiencing rapid innovation in sensing technologies, the positioning algorithms themselves, and the applicati ...

Keywords: location-based computing, pervasive computing, software architecture, ubiquitous computing

13 Distributed data sources: Exploiting online sources to accurately geocode addresses
Rahul Bakshi, Craig A. Knoblock, Snehal Thakkar
November 2004 **Proceedings of the 12th annual ACM international workshop on Geographic information systems**
Publisher: ACM Press
Full text available: pdf(1.03 MB) Additional Information: full citation, abstract, references, index terms

Many Geographic Information System (GIS) applications require the conversion of an address to geographic coordinates. This process is called geocoding. The traditional geocoding method uses a street vector data source, such as, Tigerlines, to obtain address range and coordinates of the street segment on which the given address is located. Next, an approximation technique is used to estimate the location of the given address using the address range of the selected street segment. However, this ...

Keywords: geocoder, geospatial data integration, information integration, mediator

14 Reformulating query plans for multidatabase systems
Chun-Nan Hsu, Craig A. Knoblock
December 1993 **Proceedings of the second international conference on Information and knowledge management**
Publisher: ACM Press
Full text available: pdf(1.00 MB) Additional Information: full citation, references, citings, index terms

15 Geographic base files and the world of Polk
Morton A Meyer
January 1980 **Proceedings of the ACM 1980 annual conference**
Publisher: ACM Press
Full text available: pdf(666.03 KB) Additional Information: full citation, abstract, index terms

Despite the fact that R. L. Polk & Co. is an international organization operating in western Europe, Canada and Australia as well as the United States, and has been a going and growing concern for close to 110 years, it is highly probable that most members of ACM have never heard of us. The reason, of course, is very simple. Each organization impinges only very lightly on the interests of the other. There are three divisions within Polk which rely upon computerized geographic ...

16 The Map Information Facility—a cooperative federal and private venture in geocoding
Robert T. Aangeenbrug
January 1980 **Proceedings of the ACM 1980 annual conference**
Publisher: ACM Press
Full text available: pdf(518.07 KB) Additional Information: full citation, abstract, index terms

The use of a nationwide information inquiry system based on a geographic location will be and has been heralded for some years. We take for granted that the post office, the "phone system", and a number of private firms have been and will continue to provide such service. Any of us can call most airlines and get a reasonable answer on route planning for our trips with the various options explained to us by skilled operators at anytime of day from any place. Private companies eva ...

17 Handheld computing (HHC): Extending the location API for J2ME™ to support friend finder services
David Parsons
April 2006 **Proceedings of the 2006 ACM symposium on Applied computing SAC '06**
Publisher: ACM Press
Full text available: pdf(144.52 KB) Additional Information: full citation, abstract, references, index terms

The Location API for J2ME is a standard Java mobile client API that is intended to provide a generic interface to multiple positioning technologies. Its client side object model goes beyond the provision of raw location data to enable geocoding and reverse geocoding of physical landmarks, utilising the mobile device's persistent storage. However this alone does not provide direct support for 'friend finder' type applications that encompass third party mobile devices. In this paper we propose som ...

Keywords: J2ME, friend finder, location API, location based services

18 Ubiquitous computing (UC): Route profiling: putting context to work
Anthony Harrington, Vinny Cahill
March 2004 **Proceedings of the 2004 ACM symposium on Applied computing**
Publisher: ACM Press
Full text available: pdf(232.60 KB) Additional Information: full citation, abstract, references, index terms

Intelligent Transportation Systems are characterised by a requirement for detailed information on extensive transport networks. This information is typically gathered from sensors deployed throughout the network and is used for management and maintenance operations.In this paper we present the design and prototype implementation of a context-aware route profiling application intended for use by road management authorities in the Republic of Ireland. Our design allows data from a variety of sourc ...

Keywords: ITS, context-aware, ubiquitous computing

19 Automated Generation of Visual Simulation Databases Using Remote Sensing and GIS
Martin Suter, D. Nuesch

October 1995 **Proceedings of the 6th conference on Visualization '95**
Publisher: IEEE Computer Society
Full text available: pdf(1.15 MB) Additional Information: full citation, abstract, citings
Publisher Site

This paper reports on the development of a strategy to generate databases used for real-time interactive landscape visualization. The database construction from real world data is intended to be as automated as possible. The primary sources of information are remote sensing imagery recorded by Landsat's Thematic Mapper (TM) and digital elevation models (DEM). Additional datasets (traffic networks and buildings) are added to extend the database. In a first step the TM images are geocoded and then ...

Keywords: remote sensing, geographic information systems, geographic databases, satellite images, classification, visual simulation, level of detail

20 The geographic information systems (GIS) standards infrastructure
Henry Tom
September 1994 **StandardView**, Volume 2 Issue 3
Publisher: ACM Press
Full text available: pdf(1.32 MB) Additional Information: full citation, references, index terms, review

Results 1 - 20 of 39 Result page: **1** 2 next

THE MAPEDIT SYSTEM FOR AUTOMATIC MAP DIGITIZATION

by

H. H. Holmes, D. M. Austin and W. H. Benson[1]

A system for the automatic digitization of polygon boundaries is described. Digitized map files are created from a driver tape containing identification codes and approximate centroids of polygonal boundaries (e.g., census tracts), and a film image of the map. The digitizer scans on the film plane in an automatic line-following mode, producing the first stage of the map file for the editing system. The MAPEDIT system, which can be used either interactively or in batch mode, reads maps in several standard formats and provides for combining and selecting maps by census (or other) geocodes or by longitude and latitude. This system provides several stages of data compression, analysis, and verification, including algorithms for detecting straight lines, finding corners, fitting insets of maps together and matching boundaries common to a pair of polygons. Auxiliary programs (1) provide a very high resolution (1 part in 25,000) C.R.T. plot of the map, (2) allow a detailed examination and editing of the map and (3) supply missing geocodes using auxiliary tapes such as the Medlist tapes.

[1]Lawrence Berkeley Laboratory, University of California, Berkeley, California.

SIGGRAPH '74 Boulder, Colorado

The MAPEDIT system for automatic map digitization

Full text Pdf (12 KB)

Source International Conference on Computer Graphics and Interactive Techniques archive
Proceedings of the 1st annual conference on Computer graphics and interactive techniques
table of contents
Boulder, Colorado
Pages: 59 - 59
Year of Publication: 1974

Authors H. H. Holmes University of California, Berkeley, California
D. M. Austin University of California, Berkeley, California
W. H. Benson University of California, Berkeley, California

Sponsor SIGGRAPH: ACM Special Interest Group on Computer Graphics and Interactive Techniques

Publisher ACM Press New York, NY, USA

Additional Information: abstract collaborative colleagues peer to peer

Tools and Actions: Find similar Articles Review this Article
Save this Article to a Binder Display Formats: BibTex EndNote ACM Ref

DOI Bookmark: Use this link to bookmark this Article: http://doi.acm.org/10.1145/563182.563221
What is a DOI?

↑ ABSTRACT

A system for the automatic digitization of polygon boundaries is described. Digitized map files are created from a driver tape containing identification codes and approximate centroids of polygonal boundaries (e.g., census tracts), and a film image of the map. The digitizer scans on the film plane in an automatic line-following mode, producing the first stage of the map file for the editing system. The MAPEDIT system, which can be used either interactively or in batch mode, reads maps in several standard formats and provides for combining and selecting maps by census (or other) geocodes or by longitude and latitude. This system provides several stages of data compression, analysis, and verification, including algorithms for detecting straight lines, finding corners, fitting insets of maps together and matching boundaries common to a pair of polygons. Auxiliary programs (1) provide a very high resolution (1 part in 25,000) C.R.T. plot of the map, (2) allow a detailed examination and editing of the map and (3) supply missing geocodes using auxiliary tapes such as the Medlist tapes.

↑ Collaborative Colleagues:

D. M. Austin: W. H. Benson
H. H. Holmes
P. M. Wood

W. H. Benson: D. M. Austin
H. H. Holmes

H. H. Holmes: D. M. Austin
W. H. Benson

R. D. Burris
M. K. Gleicher
D. L. Million
S. R. White

Peer to Peer - Readers of this Article have also read:

- Data structures for quadtree approximation and compression **Communications of the ACM** 28, 9
  Hanan Samet
- A hierarchical single-key-lock access control using the Chinese remainder theorem **Proceedings of the 1992 ACM/SIGAPP Symposium on Applied computing**
  Kim S. Lee , Huizhu Lu , D. D. Fisher
- The GemStone object database management system **Communications of the ACM** 34, 10
  Paul Butterworth , Allen Otis , Jacob Stein
- Putting innovation to work: adoption strategies for multimedia communication systems **Communications of the ACM** 34, 12
  Ellen Francik , Susan Ehrlich Rudman , Donna Cooper , Stephen Levine
- An intelligent component database for behavioral synthesis **Proceedings of the 27th ACM/IEEE conference on Design automation**
  Gwo-Dong Chen , Daniel D. Gajski

CARTOGRAPHIC DATA STRUCTURES: ALTERNATIVES FOR GEOGRAPHIC INFORMATION SYSTEMS

Kenneth J. Dueker
Institute of Urban and Regional Research, University of Iowa
Iowa City, Iowa 52242

Means of capturing and encoding cartographic data for machine processing is a major issue in the design and the development of geographic information systems. The available and potential technology constitutes a bewildering array of choices for system designers from which to select formats and processing capabilities to meet user applications. Designers are not provided a very precise statement of data requirements (or fidelity requirements for the data capture and processing system); therefore, the tolerable distortion rates and tolerable information losses are not well expressed so as to assess coding efficiencies. Nor is there substantial agreement as to the appropriate breadth of comparative tests, as some systems can replicate coverages well for cartographic purposes, while other systems may provide stronger analytical capabilities for processing encoded geographic data.

One of the major problems in encoding cartographic data is the lack of measures by which to assess the effectiveness of the coding. One set of effectiveness measures relates to the ability to replicate the source document in map form, while a second set of effectiveness measures relates to the use of the map or coverage data. There are a number of effectiveness measures that need developing in order to compare and test the effectiveness of alternative system approaches. A source document constructed to possess features that will test the effectiveness of alternative systems fairly and equitably needs development. This would result in the ability to establish benchmark tests by which systems could be compared.

1. INTRODUCTION

Capture and encoding of cartographic data for machine processing is a major issue in the design and development of geographic information systems. Data capture technology -- manual coding, digitizing, scanning -- and formats -- pixels, cells, grid units, points, line segments, or polygons -- constitute a bewildering array of choices for system designers. The system designer is faced with selecting data capture technology, a format that provides a capability to meet user applications, and a data processing system that is commensurate with the choice of data format and volume of data. Cartographic applications require data formats that more closely capture the fidelity of source data. Yet, some degree of spatial aggregation is necessary for inclusion in geographic information systems.

Dr. Dueker is Director of the Institute of Urban and Regional Research and Professor of Urban and Regional Planning and Geography, University of Iowa.

Support of the Energy and Environmental Systems Division, Argonne National Laboratory (ANL) is gratefully acknowledged. ANL work performed under the auspices of the Office of Land Use and Water Planning, U.S. Department of the Interior under Interagency Agreement P7434A.

There is no easy choice with respect to fidelity requirements for data capture or the appropriate level of aggregation of data. Too little is known as to the data requirements for planning and cartographic applications and the data capture and manipulation technology for large-scale applications.

A region consists of spatially varying sets of characteristics. In this paper, each set is considered a coverage, say soils, and each coverage is categorized, say into soil classes. In other words, geographic data is considered to consist of various areas of like characteristics separated by networks of lines. A single such partitioning of a region into non-overlapping zones will be referred to as a coverage (1). For example, there may be coverages showing soil characteristics, land uses, vegetation cover type, political division, or combinations of these. Linear data, such as streams, roadways, railroads, can also be considered a coverage, but with the emphasis on the network of lines rather than the bounded areas. This paper is concerned only with coverages as two dimensional objects, thus excluding pictorial representation of the three spatial dimensions as well as time-varying pictorial information, e.g., on-line character recognition. This restriction

also rules out picture processing, computer generated movies, and computer typography. Finally this paper treats the subject of coverages from a primarily problem oriented rather than technique oriented standpoint, in that the emphasis is on the relationship between encoding coverages and applications, rather than hardware/software techniques for encoding coverages.

The first section of this paper discusses geocoding options for encoding coverages. Then the data capture, formatting, storage and output of coverages is related by analogy to information theory for purposes of illuminating parallels between encoding spatial data and encoding messages for transmission, receiving and use. This analogy proves useful in that it identifies the dilemma of system designers, in that fidelity requirements for the spatial processing have not been developed and consequently effectiveness measures for the encoding of coverage data cannot be specified. Next, the paper attempts to identify the potential for error in data capture so that system designers can be alert to situations that can occur and make allowances for remedying these errors, and the development of effectiveness measures and means to compare different systems. Finally, the paper calls for the development of comparative benchmark tests that agencies could employ to evaluate vendor systems.

2. OVERLAYING COVERAGES

A coverage can be encoded as a one-dimensional representation in which the basic record is a single contiguous homogeneous tract or "polygon" coded by locating the boundary as a series of connecting points and by indicating the character or descriptor of the enclosed territory.* Ignoring any error in the drawing of boundaries or in ascribing characteristics, the accuracy of polygon encoding is limited only by the precision with which boundaries can be coded as series of connective points. One disadvantage of the polygon method occurs when two coverages have to be compared. With polygon data it is time consuming to identify the polygon in which a certain point lies, and thus its characteristics, and to compare the same point on another coverage. The advantages of organizing the data so that the character of any location can be retrieved quickly, often leads to gridding polygon data prior to overlaying or the initial adoption of a grid data structure. In a grid structure the coverage is encoded by recording the nature of each of a series of cells ordered in some conventional sequence. Compression of grid data structure is possible to eliminate a data set containing a separate entry for each cell by means of using some form of multiplier convention for sequences of repetitious cells.

With grid data it is a simple matter to compare the characteristics of a point on two coverages. The most troublesome characteristic of grid encoding is its approximation of coverages. Accuracy is directly linked to the size of the grid cell and precise replication of a coverage requires a large number of infinitesimally small cells.

Both polygon and grid cell encoding are widely used. The polygon form is adopted by systems whose major concern is data storage and accurate cartographic retrieval, and the measurement of area. The grid system is more widely used in various forms of planning, but where accuracy is less important and the ability to overlay coverages is essential, and where the range of likely demands on the system is perhaps much better defined. (1)

3. DATA AGGREGATION

A system designer must select an appropriate level of abstraction in converting a map to a coverage, and then again in converting a coverage to digital form, both as a means of reducing the sheer data volume. Level of data aggregation and geocoding options should be determined by the intended applications. Developing these interrelationships is crucial to the design of systems having multiple uses, but at the same time it is difficult to relate the designers' options to abstract data use/application categories. Yet, systems must be designed for classes of problems, not specific applications.

Data aggregation has two components, aggregation of phenomena to categories and spatial aggregation. The level of detail for coverage categories and spatial units should be compatible. A large number of categories, for say cover or soils, generates a complex coverage, which if encoded to large spatial units, such as LUNR's (2) one kilometer grid imposes a high degree of spatial aggregation. Similarly a coarser coverage classification such as MLMIS (3) utilized finer spatial units (40 acre) for predominant use assignment. Consistency of coverage categories with spatial units is crucial to selecting a geocoding option which meets intended data uses/applications.

It is essential to recognize that comparison of systems is made difficult if they have been designed for different applications. For example, a system designed principally for cartographic uses may more exactly replicate coverages, but may be incapable of extended applications, while a system less capable of replicating coverages may have more flexible geocoding structures that enable more powerful analytical processing, such as overlaying or generating slope maps.

4. A COMMUNICATIONS SYSTEM FOR GEOGRAPHIC DATA

A generalized communication system is comprised of elements: a data source from which a message is encoded, a transmitter from which a signal is emitted, a channel for communicating the signal, a receiver for receiving the signal and converting it back to a message to the final destination. This problem of sending and receiving messages

---

***Polygon encodings can be generated by direct digitizing of zones, which requires editing by reconciling line segments of adjacent polygons to eliminate slivers and overlaps; or encoding lines and center points and creating polygon records. Both methods require manipulation to generate clean coverages in a polygon format.**

through the use of the system which is constrained by channel capacity and the presence of perturbances (noise and distortion) is the general case from which this discussion of encoding geographic data is an application.

A geographic information system also has elements of a sender, receiver, message, signal and channel. The problems of channel capacity and transmission cost may be considered as analogous to computer storage size and machine processing cost. Of particular interest in this paper is the encoding problem. In the generalized communication system, information theory is used to measure the amount of information (in units called "bits") that is contained in the data being processed, and this theory aids in the evaluation of alternative encoding schemes to eliminate redundancy through efficient coding. Information theory is employed here to consider means of reducing redundancy in the storage and processing of geographic data. The adaptation of the generalized communication system model to geographic data demonstrates that the terminology or organization concepts for information theory is useful in the analysis of geographic data handling problems, particularly in assessing the efficiency of encoding data.

In applying information theory, a digital coverage consisting of a quantized arbitrary matrix is considered a set of messages. The gray level of intensity for each cell is a "message." If there are m gray levels or intensity categories, the total amount of information in an n-by-n digital coverage (which is average amount per element times the number of elements) can be as high as $n2\log_2 m$ bits. The actual information content depends upon the probabilities with which the gray levels occur. Physical pictorial media can be used to store information at extremely high densities. However, pictures encountered in practice (television images, line drawings, printed pages, etc.) have information content that fall appreciably short of their potential capacities, often by a factor of two or more. The difference between potential and actual information content is called redundancy. (4) Efficient encoding of pictures or coverages is possible if redundancy is minimized, and there has been considerable effort made in devising coding schemes to represent a picture or coverage as compactly as possible.

This process of approximating the picture acceptably (where the standards of acceptability may be either objective or subjective) by another picture that has lower information content has been directed largely toward the goal of television bandwidth compression. However, the manual abstraction of coverages from pictures and other sources provides an initial abstraction of the information from which further approximation is possible. The object is to reduce redundancy and not content.

The process of creating coverages from images or pictures is the first step in efficient coding of pictorial information. Cells within a single map segment of a coverage by definition have the same message or value. Hence, within a map segment a message of the next cell is a predetermined value rather than a probability of that element being the same as the preceding element message. Thus, the source document is an abstraction of reality, which when encoded enables the reduction of redundancy. Consequently, encoding of coverages reduces to a case of coding long "runs" of repeated message. Therefore, it is then economical to encode the first message of each run and then the length of the run rather than encoding each message or cell in a sequence, and all of the detail that appears in a picture is replaced by a simpler coverage that looks like the original but that has a lower redundancy. The degree of "compression" that can be obtained by approximation methods for generating coverages is generally greater than that obtainable by encoding techniques alone.

There are two basic methods used to approximate pictures; these are sampling and quantization. Sampling consists of taking values at a finite set of points, and approximates the surface by interpolating analytically simple functions through these values. In quantization, one allows the function or picture to take on only a finite set of values or quantization levels, (replacing the actual value at each point by the quantization level closest to it).

In approximating a function or surface from sampling methods one can sample from equally spaced points or a rectangular array, though it is sometimes desirable or necessary to use unequally spaced points. Contouring routines such as SYMAP or trend surface analysis, using polynomial interpolation, can be used to approximate the value at any point in an n-by-n matrix representing an n-by-n digital picture.*

Figure 1 represents efficiencies in coding geographic data. Figure 1(a) represents an n-by-n matrix of fine cells or pixels with m gray tone levels. With no predictable pattern the information content is $n2\log_2 m$ bits. A complex urban land use scene would have a lower level of randomness or entropy and a rural land use scene would even be more orderly or less complex a pattern. Regular sampling to approximate the image implies a larger cell size (see Figure 1(b) where $n > n'$) and quantization (where $m > m'$), imposes a further ordering to the image.

Figure 1(c) suggests boundaries are drawn around contiguous sample cells of like category, thereby creating the coverage. Figure 1(d) represents one type of encoding where row i consists of n'_1 columns of $m'1$, n'_2 columns of m'_2, and $n'3$ columns of m'_3. In the case of (b) the information loss can be estimated and in (d) the coding efficiency can be estimated. (Run length coding, as this row record with a multiple for sequences of repetitious cells is called, is only one type of encoding. Polygon or line segment encodings are often used. The

***Most grid systems record the predominant use, primary and secondary categories, or actually measure the area of each quantization level within the grid. Nevertheless, grid units are a means of spatial sampling and only the attributes of the sample points differ.**

run length coding illustrates best for the purpose of this discussion the potential for compression or efficiency of encoding.)

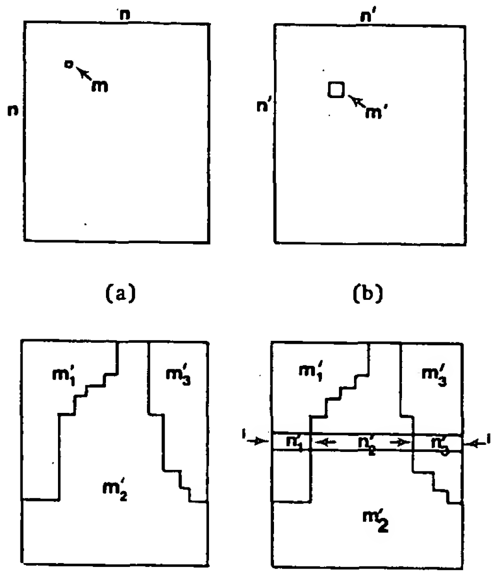


Figure 1
Information Content and Efficiency of Coding

The utility of information theory is that it forces system designers to ask important questions about information loss, redundancy, and coding efficiencies.

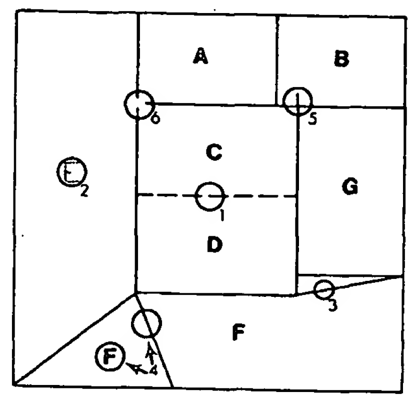
5. ERRORS IN CAPTURE AND ENCODING

The capture, encoding and processing of coverages provides considerable "opportunity" for the intro duction of error. Although source document errors are not within the scope of this discussion several of the more basic source document errors will be identified. Most attention is given to the encoding errors that occur in the capture of coverages. Finally, some attention is given to logical errors associated with the processing of data, particularly that processing associated with editing the file to insure completeness and error detection.

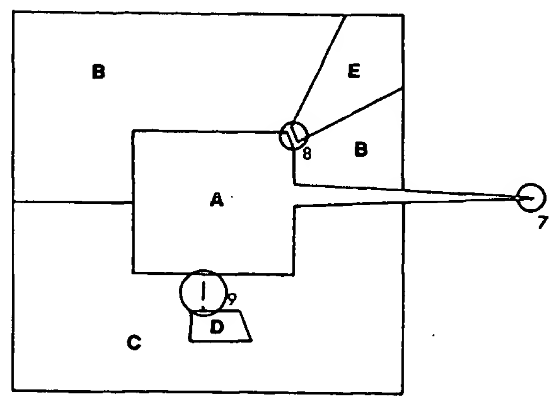
Source document errors that should be removed prior to data capture, but which if not, should be detected for correction during or following digitizing. Source document errors usually consist of missing line segments, missing center identifiers, redundant center identifiers, or redundant line segments.

Although it would be desirable to assume error free source documents, it is unlikely that error free source documents can be achieved. Consequently, the data capture and processing system must be capable of detecting and identifying the errors for subsequent remedy.

Figure 2 identifies various kinds of encoding errors which can occur because of or in spite of error free source documents. Encoding error type one is a failure to encode a line segment; type two, a failure to encode a center; type three, a failure to end a line segment; and type four, a redundant encoding of a center or a segment.



(1) Failure to encode a line segment,
(2) A center,
(3) Failure to end line segment
(4) Redundant encoding of center or segment
(5) Overshoot,
(6) Undershoot at junction points



(7) Digitizer error
(8) Inability to maintain a narrow isthmus
(9) Failure to identify contained and containing polygons

Figure 2. Encoding Errors

Source: Goodchild (5) (modified)

Encoding errors types five and six consist of the overshoot and undershoot problem resulting from difficulties of digitizing line segments to meet at a common junction point.

Encoding errors specifically with digitizing are represented by encoding errors seven, eight and nine. Error type seven is a pure digitizer error caused by the movement to another point than intended; type eight is where a narrow isthmus was not maintained and the digitizer created two polygons where one was intended, and error type nine consists of failure to identify contained and containing polygons.

Source document errors and/or encoding errors that escape detection show up as logical errors in machine processing. These must be detected and corrected. Two types of edits are employed to purge files of errors. One type are those which compare points which should coincide but which do not because they were digitized separately; similarly, separately encoded lines and areas may be compared for reconciliation. In this way undershoots, overshoots, and slivers are removed. These edits can contribute to error if the threshold for closing gaps eliminates an intended isthmus or short line segment. Graph Theory edits to purge files of missing or superfluous line segments and center identifiers are often employed. These consist of chaining around polygons and junctions to ensure polygons have a single identifier and that line segments leading to junctions close correctly.

Although this discussion of capture and encoding error pertains directly to digitizers, a scanner which detects the presence of cells that are part of a line segment and converts these line segments to polygon records must deal with the same problem, as well as the additional problem of converting cells that indicate the presence of a line segment to polygon records. Alternatively, scanning the coverage to create grid data is a problem of filling the cells with the appropriate polygon center identifiers.

Assuming a clean source document, scanning for the presence and absence of line segments consists of smoothing small scan cells into line segments, identifying junctions, and describing polygons and associating the correct center identifiers with each polygon. Such is done in the CGIS system. (6)

Direct scan to grid units is an aggregation problem; one of combining small scan cells into grid units for storage and use. ORRMIS (7) performed separate scans for each category or classification of a coverage. For coverages consisting of land use or soil type, a manageable number of scans enabled encoding a coverage, whereas a large number of uniquely identified areas requires an inordinate number of scans to encode a coverage for say census tract identifiers.

Scanning or automatic line following technology is more likely to encounter difficulty with source document errors, whereas a digitizer operator can recognize and correct many source document errors. Consequently, line gaps, uneven line widths, intensity variations of patterns on source documents, may cause considerable problems with fully automated data capture systems. In addition, the logic of creating line segments and polygons from scan cells may create error situations.

Whether data capture is accomplished by wholly manual methods, such as overlaying a grid on a map; by man-machine interaction, such as digitizing; or completely by machine as in scanning, the potential for error must be anticipated and procedures developed to insure an adequate level of quality of the encoded data to meet the purposes of the intended application.

In digitizing, the degree to which interactive editing is employed may be a function of the level of digitizer operators. High level operators may be capable of judgments to interactively edit data; whereas if lower level digitizer operators are employed, post edit of their work may be more advisable.

6. EFFECTIVENESS MEASURES AND BENCHMARK TESTS

One of the major problems in encoding geographic data is the lack of measures by which to assess the effectiveness of the encoding. This section of the paper attempts to identify possible effectiveness measures that would allow comparison of data capture technology and encoding methods. On one hand, there are a set of effectiveness measures that relate to the ability to replicate the source documents in map form and there is a second set of effectiveness measures that are more user oriented. These latter measures relate to the marginal utility of additional precision of data with respect to decision making, which at this time can only be approached by setting standards or requirements related to the degree of aggregation necessary for different classes of problems or applications.

With respect to the former problem, that of replicating coverages or overlays of coverages, several effectiveness measures are suggested. If a coverage is assumed to consist of a large number of pixels or scan cells, an effectiveness measure would be the proportion of pixels that are correctly classified through the polygon and grid coding. Similarly, an overlay of two or more coverages could also be assessed by calculating the number or percent of pixels correctly classified. With respect to area measurement, the area estimates from polygon and grid coding could be compared to the pixel count, to estimate the percent error in area measurement. Similarly, although beyond the scope of this investigation, effectiveness measures are generated from cartographic or map standards for cartographic applications of the coded data.

In addition, there need to be comparative measures developed with respect to quality control that encompass the error rate from the source document to the capture to the editing. What is needed here are per record, per cell, per polygon, per frame, per hour error rates for different system types by coverage type. One such measure

would be to regress digitized points for a line of a known function. Finally, there need to be effectiveness measures that relate to the applications of updating, edge matching and retrieving data.

Benchmark tests enable comparison of alternative systems. With respect to geographic data encoding, one objective is to compare the accuracy of the final product with the original input. This kind of benchmark test checks the efficiency of the hardware system in the functioning of the software programs. Another objective is to compare the accuracy of area measurement, especially after overlaying coverages. This requires use of standardized input in the comparison of the outputs utilizing statistical tests. Yet, several questions remain:

1. How broad or how narrow should all encoding tests be? Should it only test the cartographic replication of source documents? Should it include overlay analysis? Should it include other analysis, e.g., generating slope maps?

2. Can one general benchmark test be constructed for both digitizers and scanners?

3. Should there be separate benchmark tests for evaluating source documents? For encoding errors? For logical errors?

7. IN CONCLUSION

This paper has identified a series of issues in encoding coverages which will require attention within the next few years. The major unresolved issue is the breadth of the tests to compare encoding processes. Geographic data handling for statewide land use applications requires addressing the issues raised in this paper. Although systems will have to be designed before these issues are resolved fully, the issue identification process alerts system designers to potential problem areas.

Until some of these issues are resolved, system designers should caution designers of geographic information systems of the potential for errors and delays and cost overruns when attempting to encode and replicate a large number of complex coverages. Presently a coarser statewide system is more appropriate, while at the same time undertaking prototype developments in smaller study areas to test more sophisticated encoding techniques and to develop staff capabilities.

REFERENCES

(1) Michael F. Goodchild, Error estimation for geographic data structures, Department of Geography, University of Western Ontario, 1974.

(2) R.L. Shelton and Ta Liang, Land inventory systems: The Cornell experience, Ithaca: Cornell University, 1973.

(3) Land management information in Minnesota: The beginning of a statewide system, Minneapolis: Center for Urban and Regional Affairs, University of Minnesota and Saint Paul: Minnesota State Planning Agency, 1972.

(4) Azriel Rosenfeld, Picture processing by computer, New York: Academic Press, 1969.

(5) Michael F. Goodchild, PLUS: Technical user's guide, London, Ontario: Department of Geography, University of Western Ontario, 1974.

(6) R.F. Tomlinson, An introduction to the geographic information system of Canada land inventory, Geographic Information Systems ASP-ACSM Annual Convention, Washington, D.C., 1967.

(7) R.C. Durfee, ORRMIS Oak Ridge regional modeling information systems, Oak Ridge, Tennessee: Oak Ridge National Laboratory, ORNL-NSF-EP-73, 1974.

Cartographic data structures: alternatives for geographic information systems

Full text Pdf (81 KB)

Source **International Conference on Computer Graphics and Interactive Techniques** archive
Proceedings of the 3rd annual conference on Computer graphics and interactive techniques
table of contents
Philadelphia, Pennsylvania
Pages: 167 - 172
Year of Publication: 1976
ISSN:0097-8930
Also published in ...

Author Kenneth J. Dueker University of Iowa, Iowa City, Iowa

Sponsor SIGGRAPH: ACM Special Interest Group on Computer Graphics and Interactive Techniques

Publisher ACM Press New York, NY, USA

DOI Bookmark: Use this link to bookmark this Article: http://doi.acm.org/10.1145/563274.563306
What is a DOI?

↑ ABSTRACT

Means of capturing and encoding cartographic data for machine processing is a major issue in the design and the development of geographic information systems. The available and potential technology constitutes a bewildering array of choices for system designers from which to select formats and processing capabilities to meet user applications. Designers are not provided a very precise statement of data requirements (or fidelity requirements for the data capture and processing system); therefore, the tolerable distortion rates and tolerable information losses are not well expressed so as to assess coding efficiencies. Nor is there substantial agreement as to the appropriate breadth of comparative tests, as some systems can replicate coverages well for cartographic purposes, while other systems may provide stronger analytical capabilities for processing encoded geographic data.One of the major problems in encoding cartographic data is the lack of measures by which to assess the effectiveness of the coding. One set of effectiveness measures relates to the ability to replicate the source document in map form, while a second set of effectiveness measures relates to the use of the map or coverage data. There are a number of effectiveness measures that need developing in order to compare and test the effectiveness of alternative system approaches. A source document constructed to possess features that will test the effectiveness of alternative systems fairly and equitably needs development. This would result in the ability to establish benchmark tests by which systems could be compared.

↑ REFERENCES

1 Michael F. Goodchild, Error estimation for geographic data structures, Department of Geography, University of Western Ontario, 1974.

2 R. L. Shelton and Ta Liang, Land inventory systems: The Cornell experience, Ithaca: Cornell University, 1973.

3 Land management information in Minnesota: The beginning of a statewide system, Minneapolis: Center for Urban and Regional Affairs, University of Minnesota and Saint Paul: Minnesota State Planning Agency, 1972.

4 Azriel Rosenfeld, Picture processing by computer, New York: Academic Press, 1969.

5 Michael F. Goodchild, PLUS: Technical user's guide, London, Ontario: Department of Geography, University of Western Ontario, 1974.

6 R. F. Tomlinson, An introduction to the geographic information system of Canada land inventory, Geographic Information Systems ASP-ACSM Annual Convention, Washington, D.C., 1967.

7 R. C. Durfee, ORRMIS Oak Ridge regional modeling information systems, Oak Ridge, Tennessee: Oak Ridge National Laboratory, ORNL-NSF-EP-73, 1974.

↑ CITINGS

John F. McIntosh, The interactive digitizing of polygons and the processing of polygons in a relational database, Proceedings of the 5th annual conference on Computer graphics and interactive techniques, p.60-63, August 23-25, 1978

↑ Peer to Peer - Readers of this Article have also read:

- Data structures for quadtree approximation and compression **Communications of the ACM** 28, 9
  Hanan Samet
- A hierarchical single-key-lock access control using the Chinese remainder theorem **Proceedings of the 1992 ACM/SIGAPP Symposium on Applied computing**
  Kim S. Lee , Huizhu Lu , D. D. Fisher
- The GemStone object database management system **Communications of the ACM** 34, 10
  Paul Butterworth , Allen Otis , Jacob Stein
- Putting innovation to work: adoption strategies for multimedia communication systems **Communications of the ACM** 34, 12
  Ellen Francik , Susan Ehrlich Rudman , Donna Cooper , Stephen Levine
- An intelligent component database for behavioral synthesis **Proceedings of the 27th ACM/IEEE conference on Design automation**
  Gwo-Dong Chen , Daniel D. Gajski

↑ This Article has also been published in:

- **ACM SIGGRAPH Computer Graphics**
  Volume 10 , Issue 2 Summer 1976

MapMarker 4.1.0

Support Notes

August, 1998

MapInfo Corporation
Troy, New York

Introduction

This document is to be used as a reference for the MapMarker® 4.1.0 product. It describes the system requirements, provides additional product information, and explains known problems with the product. This document should be used in conjunction with the MapMarker 4.0 *User's Guide* and the 4.1 Supplement.

System Requirements

MapMarker 4.1.0 runs on the Microsoft Windows® 95 operating system or Microsoft Windows NT® 3.51 or later.

The minimum system requirements for MapMarker 4.1.0 are a 486 with 8 MB of memory for Windows 95 and 16 MB of memory for Windows NT. The recommended minimum is 16 MB of memory for Windows 95 and 32 MB of memory for Windows NT.

Product Information

What's New in MapMarker 4.x?

Candidate Visualization

Candidate visualization allows you to see where potential matches fall on a map before you make your choice. This feature is accessible via a Map button in the interactive dialog and the new Quick Find dialog. You can now select the point on the map that represents your match choice and MapMarker will geocode to that record. The Quick Find feature allows you to view the candidates on a map, but does not return the information to your table. Quick Find is a quick way to confirm an address. Candidate visualization uses StreetWorks or StreetInfo tables as the background street network on which the candidates are placed. StreetWorks ships with MapMarker.

Attribution

This feature allows the user to attach data from another table to geocoded records. Any table in MapInfo format is suitable for this process– boundary and point files, demographic tables, or non-geographic tables. Any information that is stored in the attribution table can be attached to the record in your database when MapMarker makes a successful match. The user needs only a common link between the record in the geocoding table and the attribution table.

Attribution can be performed either when geocoding the table or as a batch process separate from a geocoding pass. Batch attribution is faster than geocoding/attribution, and attribution using a column-to-column match is faster than using a geographic join.

Quick Find

MapMarker 4.x's new search feature allows you to type in a single address record to search. MapMarker will return the complete address if it makes a match. (Note: this feature does not geocode the record.) Additionally, if there is more than one candidate, each one can be mapped to help decide the best match. The Quick Find feature is under the Search menu.

Geographic Precision

The precision of the coordinates returned by MapMarker has been increased to five significant digits. This increases the positional accuracy at which geocoded records display on a map. If you set MapMarker's street offset at 0 feet for displaying over a StreetWorks street table, the average positional error from the center of the street would be plus or minus 3 feet.

SpatialWare 2.2 Support

MapMarker supports the spatial data type SW_GEOMETRY when geocoding remote tables. Users of SpatialWare 2.2 for Oracle can geocode with MapMarker, and store resulting coordinate information and spatial objects directly in the remote table. MapMarker also continues its support of spatial data in X, Y columns for remote tables.

Quick Geocode

For geocoding without all the setup, the Quick Geocode button allows you to click and go. MapMarker will geocode the table with the current preferences without displaying the Geocode dialog. This feature requires that the table be opened previously in MapMarker in order to set the columns and geocoding preferences.

The Quick Geocode feature does not work for remote tables because the feature uses the table's metadata to obtain the geocoding preferences. Metadata cannot be stored in remote tables.

CASS

As in previous 3.x versions, MapMarker version 4.1 meets the USPS CASS requirements for address standardization, including the ability to append ZIP+ 4 information to your data.

Table Modify

You may now change the structure of your table once it is open in MapMarker. This is useful when you want MapMarker to return additional information from the Address Dictionary, but you did not set up a column for it before opening the table. Instead of altering the structure in MapInfo Professional, you can keep the table open in MapMarker and make the changes there. These changes include adding or removing columns, renaming columns or changing their type.

MapMarker Address Dictionary

The Version 4.1.0 Address Dictionary for MapMarker has been updated with recent information from three sources:

- U.S. Postal Service address and ZIP+4 information (vintage June 1998)
- Street data from the U.S. Census Bureau TIGER 95 files (release date September 1996)
- ZIP+4 Centroids from GDT (vintage April 1998)

This address dictionary contains a wealth of information related to addresses and street information. It is updated bimonthly as required by CASS. Even if you are not geocoding to CASS standards, you can be sure your records are being geocoded to the best match possible.

Helpful Hints & Known Problems

Geocoding from CD

When geocoding with data on CD and a second pass is needed with CD #2, MapMarker may give an error stating an address file may be corrupt. The workaround is to go to System Preferences, select the Dictionary tab, and click OK. This will re-initialize MapMarker with the second CD.

Input/Output Selection Dialog

With the addition of more output information (e.g., Firm output names, Apartment numbers, suites, etc.), recommended output lengths for Street and Firm names aren't always long enough. If the output result for a Street or Firm is too short, it will be truncated in the respective output field.

If a field name is too long to be seen in the input or output edit boxes, hold the cursor over the edit box. MapMarker will show a popup box with the full name of the field.

Modify Table

MapMarker 4.x does not support modifying MapInfo Version 410 tables. Opening a Microsoft Access table natively in MapInfo Professional 4.x makes it a version 410 table. To add columns to Version 410 tables, please open the table in MapInfo Professional.

Addresses

If a place has an address such as "Lenz and Riecker," MapMarker will attempt to geocode it as an intersection. The workaround is to add "Plz" or "Place" at the end of the address. MapMarker then treats the entire string as one street or place and properly suggests "Lenz and Riecker" as a match candidate. (Bug #3678)

Attribution

In order to use the Batch Add Attributes feature, the table must be geocoded at least once.

Quick Geocode

A table must be geocoded once for the Quick Geocode button/menu to be enabled.

Labels in the Candidate Map are zoom layered at 10 miles. This means that the labels are visible at a zoom level of 1 to 10 miles. If labels are added at a 1-mile zoom, and the user zooms out from the map, the visibility of the labels will be turned off when the zoom level reaches 10 miles.

If the map is too cluttered when viewing candidates in Quick Find, try removing the following point layers from the Maps dialog (System Preferences > Maps): Points Cultural, Points Natural, and Area Landmarks.

When using the Browse feature, any part of the street name can be used to search the User Dictionary, e.g., entering 2 G, troy, NY, 12180 returns Garden CT, Garden Way, etc.; entering GL returns Glen Dr, Glenkill, etc.; entering GLO returns Global View only.

DBF

If for some reason a dbf file has a time stamp dated sometime in the future, you will get a "Can't Create Tab File" error message when you try to open the file. The fix is to open it in dBase. The time stamp will correct itself. (Bug #3679)

Geocode Dialog

If you check Specify Log File Name in the Log File page of the Geocode dialog, a path may be entered as well as a name for the log file (e.g., C:\temp\Us_addr.log). The default is to place the log file in the same directory where MapMarker is installed. If only Update Log file is selected, MapMarker will update the MapMarkr.log in the directory where MapMarker is installed.

Installation/Uninstallation

When installing MapMarker as a network Server install to a Novell Network the directory name must be 8 characters or less for the client setup to properly run. If, when running the client setup program from the network, the installer suddenly disappears, check the length of the main directory name.

If you have problems installing MapMarker (e.g., you get an error message that a directory does not exist), check that your Windows, Windows System (System32 on NT), Temp directories, and Windows Temp directories are not set as read only.

When installing just the "ODBC" portion of MapMarker (say, over a previous install of everything else), there are no shortcuts added for the ODBC Installer or ODBC Driver Help. The workaround for running the ODBC Installer is to run the Setup.exe in the \ODBC subdirectory of the MapMarker install directory. (Bug # 3729)

MapMarker Client/Server Toolkit

When the MapMarker server is set to be an automatic service on startup, it is unavailable until a message displays telling the user that a service did not start up properly. After that, the MapMarker server starts properly. To work around the problem, do the following:

1. Go to the Control Panel.
2. Choose Network > Services > RPC Configuration> Properties.
3. Change the Name Service Provider from Windows NT Locator to DCE Cell Directory Service and enter your IP address in the Network Address edit box.

Note: It seems that anything typed into the Network Address edit box corrects the problem if you are not connected to a network nor have TCP/IP. (Bug # 17789)

In Visual Basic 5 only, the MapMarker OCX does not tab through the input fields. It works fine in Visual Basic 4. (Bug # 3742)

When a user who has administrator rights installs MapMarker on NT, users who then log on do not have access to the MapMarker service in the Control Panel. (Bug # 3738)

If the OCX returns the message "Bad Input Address," the error code is not set to 14. If you invoke GeocodeAddress() or GeocodePostalCentroid() via OLE Automation, the error code is returned correctly.

When using the MapMarker OCX/Server, a number of conditions will generate "Bad Input Address" error messages in the status information field. The list that follows describes the possible cases and indicates whether they occur at the OCX level or at the Geo-Engine level.

At OCX Level

(1) The address and ZIP Code fields are both empty.

(2) The city/state pair and ZIP Code are both empty.

(3) The city or state is misspelled or an invalid ZIP Code is supplied.

To avoid this error condition supply one of the following:

(a) A ZIP Code (only)

(b) Street and ZIP

(c) Street and city/state pair

(d) Street, city/state pair and ZIP

The user may have simply misspelled city or state, or used the wrong ZIP Code.

At Geo-Engine Level

The "Bad Input Address" message may still be returned even if you supply a valid street address and a valid city/state combination but omit the ZIP Code.

This occurs when a match on ZIP Code is required, but a ZIP Code has not been entered. Check the MapMarker Server geocoding parameters. The parameters are defined in the Registry.

Users can use the registry editor (regedit.exe (W95) **or** regedt32.exe (NT)) to find the settings under key:

"HKEY_LOCAL_MACHINE\SOFTWARE\MapInfo\MapMarker\4.0\GEOCODING\" and

"HKEY_CURRENT_USER\SOFTWARE\MapInfo\MapMarker\4.0\GEOCODING\"

Change the RequireZipCode parameter from 1 to 0, then restart the mm_serve again.

The OCX will be chopped off if the user specified Small Fonts in the System Settings. It works with Large Fonts. (Bug # 3756)

If an Administrator installs MapMarker and the MapMarker service is started, a user who logs onto the same machine is not able to hit the service via the OCX. The service is running as viewed through the Control Panel, but the user rights do not seem to give them access to the service. A user can run the service as a console application, but an Admin installation locks out users who then log on. (Bug #4592)

ZIP Codes

MapMarker will not provide latitude/longitude information for some ZIP Codes if the information is unavailable in the source data. (Bug # 3675)

ODBC/OLEDB

MapMarker will give an error when only 1 record exists on an input table. The reported error comes from the Microsoft OLE DB provider. It does not occur when you have 0, or 2 or more records. The error is "The rowset cannot scroll backwards."

When geocoding Excel tables make sure the table doesn't contain any apostrophes ('). This causes the geocoding operation to stop at that record.

Intersolv's IUS driver will not work in conjunction with an Abstract Data Type and OLE DB.

Intersolv's Excel driver is unsupported.

If an input table is not required to have a primary/unique key, it may have duplicate values. The output table **must** have a key. When the insert is performed an error will be produced upon updating with a duplicate key (bug #4585).

MapMarker 4.x does not support geocoding Excel tables larger than approx. 65,000 records.

After geocoding a small (less than 60 record) Excel table MapMarker may not update the browser. The Workaround is to close the table and reopen it.

The Remote Table property page may be available when geocoding a single table, but it should be ignored because it only works when geocoding to an output table.

If output columns are set to numeric field types where they should be characters or vice versa, then the table is not updated with geocoding results. See the "Output Columns" section of the MapMarker documentation for the output column types and widths required for each output field. (Bug # 3753)

In the performance tab of the ODBC Sybase Data Source Setup dialog, be sure to set the Prepare Method to 2 – Full, and the Select Method to 1 – Direct. The default settings of None and Cursor may cause locking if there are concurrent users accessing the database.

Be sure that you set your rollback segment or temp space to a size large enough to accommodate the size of your database.

MapMarker is not able to geocode view tables.

If you install just the ODBC components (over a previous installation of everything but ODBC), there are no shortcuts added to the MapInfo program group for the ODBC Installer or the ODBC Driver Help. (Bug # 3736)

For a table created via SQL, **do not** insert/update value(s) into a numeric primary key or unique index column that have values larger than the field's numeric precision. MapMarker will not geocode records created in this manner.

Example:

Microsoft Access has a Numeric Double with a precision of 15.

Unacceptable Input	Acceptable Input
2222222222222222215	222222150000000000

Longitude and latitude columns must allow for NULL values.

If another user deletes rows from your remote table after MapMarker opens the table, MapMarker will report the following error when you try to view one of the deleted records in the MapMarker browser: "You must close and then reopen the table to receive a correct view and record count." If records have been deleted, modified, or added, you may wish to re-geocode the table to obtain an updated MapMarker log file, which contains an accurate count of the number of geocoded records, and to geocode any records that were added to the table.

If a remote table contains less than 20 records, MapMarker will not update the browser after geocoding. The workaround is to close the table and reopen it.

If a remote table is opened, and its name is more than 12 characters long, only the **last** 12 characters are displayed in the Table box of the Select Input Columns dialog. (Bug # 3709)

MapMarker requires that each remote database table that is to be geocoded have a unique index or primary key. The following matrix outlines the index and primary key data types that are supported in MapMarker v. 4.1.0.

Supported Index and Primary Key Types for Geocoding Remote Tables

Microsoft SQL Server	Oracle	Microsoft Access
		varchar
tinyint	char	byte
char	number	char
int	varchar	long
smallint	Varchar2	counter
varchar		short
		single

Note: Although MapMarker will geocode an Access table that is indexed on a field type of double or geocode a SQL Server table that is indexed on a field type of float, these fields are not supported. MapMarker may not write the geocoding results or may write the geocoding results to the wrong record in these cases.

When setting up output columns to store longitude and latitude coordinates, be sure to "match" what is specified for these columns in the table's Map Catalog. MapMarker does not automatically choose the long/lat columns contained in the Map Catalog. Specifically, the output longitude and latitude columns selected by the user must agree with the longitude and latitude columns specified in the Map Catalog.

If MapMarker determines there is a datum conflict between what is set in System Preferences and what is listed in a remote table's Map Catalog, MapMarker prompts you for a decision. If you say yes, the MapCatalog will be changed to match the System Preference setting (either NAD83 or NAD27). If you say No, MapMarker will geocode the table to NAD83 but it will not alter the Map Catalog. When you display the geocoded points in a Map window using the original projection that is listed in the Map Catalog, the points with NAD 83 coordinates may display at the wrong locations. To get around this problem, modify the MapInfo Map Catalog (either via MapInfo Professional or another SQL utility) and switch the Projection to NAD 83 manually. A catalog entry of NAD 83 would display in one of the following two ways: Earth Projection 1, 33 **or** Earth Projection 1, 74.

Upgrading from 3.x

When you are upgrading from MapMarker 3.x to 4.x, and you keep both versions on the same machine, the entry in the Add/Remove programs for MapMarker 3.x now references MapMarker 4.0. To remove MapMarker 3.x, please use the **uninstal.exe** program from the directory where MapMarker 3.x resides.

User Dictionaries

To initialize the MapMarker Server with User Dictionaries you must enter the path in the following location in the Registry:

HKEY_LOCAL_MACHINE\Software\MapInfo\MapMarker\4.0\System\

UserDictionaryPath

To run the registry editor, click the Start button in Win 95/NT 4.0. Choose run, type regedit, and click OK. Browse to the above location and double-click on the UserDictionaryPath string under System. **Note:** It is highly recommended that you back up your registry before making any changes. (Bug # 5152)

When creating a User Dictionary, make sure all the information in the City fields is capped. If it is not, MapMarker may geocode improperly.

When creating a User Dictionary make sure the MapInfo table's projection is NAD27 **or** NAD83 before moving to the Create User Dictionary Wizard.

When Selecting the State fields in Step 2 of 3 make sure that the field has the two-letter abbreviation for the particular state. The 2-digit FIPS code will cause problems when geocoding with the User Dictionary in question.

Batch Files

When creating a batch file under Windows 95/98, the batch file must be edited in a text editor to add a line so that MapMarker processes the first file before proceeding to the next. Refer to page 7 of the MapMarker 4.1 Supplement for more information.

Log File

If a geocoding log from a previous geocoding session has not been closed, you will not be able to reopen the table in MapMarker. A message displays saying that you may not have read/write access to the table. The table will not open until the MapMarker log file is closed from NotePad or whichever text editor you are using (Windows 95 only).

CD Browser

If the CD browser is open from CD #1 then CD # 2 is placed in the CD drive the operating system will ask that CD #1 be placed back in the CD drive.

API

Using GeoEngGetStatesLicensed() and GeoEngGetStatesFound() in Microsoft Visual Basic may cause abnormal termination of your application.

API/OLE Automation Changes

MapMarker OLE Automation Function Call Changes

MapMarker 3.x	**MapMarker 4.x**	**Comment**
GetLastErrorCode()	LastErrorCode	Removed GetLastErrorCode method replaced with LastErrorCode property
GetStringBinding()	StringBinding	Removed GetStringBinding() method replaced with StringBinding property

MapMarker OLE Automation Function Call Additions

MapMarker 4.x	Comment
DatabaseTypes	Property that shows the available databases bit flag: 1-street, 2-zip, 4-User.
GeocodeCheckDbAvailability()	Checks for available databases.
GeocodeGetServerVersion()	Retrieves server version number.
GeocodeGetStatesFound()	Gets a list of States found in the dictionary path.
GeocodeGetStatesLicensed()	Gets a list of States licensed.
GetCandidateFirmAt()	Retrieves the Firm of the specified candidate string.
GetVersionNum()	Gets the OCX version number.

MapMarker API Function Call Additions

MapMarker 4.x	Comment
GeoEngGetStatesFound()	Gets a list of States found in the dictionary path.
GeoEngGetStatesLicensed()	Gets a list of States licensed.
GeoEngGetVersion()	Retrieves the version of the GeoEngine.

Miscellaneous

Performance

To improve performance, sort your database by ZIP Code. When you do this, MapMarker can decrease geocoding times by up to 40%, depending on geocoding preferences, database size, and the location of addresses in the database.

dmti™
Spatial

White Paper

Sept. 1998 - MARCH 1999

Street Intersection Data Model For FSA Boundaries

The Next Generation of Postal Code Boundaries

really smart spatial solutions™

www.dmtispatial.com

This White Paper is intended to familiarize you with the newly developed Street Intersection Data Model for Forward Sortation Area (FSA)* Boundaries.

First it is important to explain the background and purpose of FSA Boundaries.

Background on FSAs

Forward Sortation Areas (FSAs) are polygons representing the boundaries that encompass postal code points with the first three characters in common, designating a postal delivery area. A postal code is comprised of an FSA LDU* or in other words a Forward Sortation Area and a Local Delivery Unit. For example, both postal codes V5L 2H2 and V5L 2H3 would be found in the same FSA - "V5L".

FSA Boundaries do not typically follow other boundaries such as municipal or census boundaries; they are unique unto themselves.

The FSA Boundaries are polygons that encompass the six digit postal code points which start with the same FSA designation in common, and conform to the streets and other physical features where applicable. For example all postal codes starting with the FSA designation of "V5L" will be contained within the same FSA Boundary and all those starting with "V5N" would be contained in a separate FSA Boundary.

FSAs are constantly being created and updated as new areas or regions are being developed or amalgamated, for example, the development of new subdivisions or the annexation of surrounding regions.

Geocoding*, the ability to provide geographic coordinates for an address so that it can accurately be placed on a map, is used to determine the level of accuracy for the postal codes and their boundaries. Postal codes are rated according to how they were geocoded. A database is geocoded first to block-face* (street level) then to urban and rural Enumeration Area* (EA) centroid*.

Purpose of FSA Boundaries

There are many possible uses for FSA Boundaries.

A sales and marketing manager may assign sales territories based on assigning a combination of FSAs to their sales representatives. Customer sales may then be geocoded by FSA to the map for a thematic representation of sales volume for example. By assigning sales territories by FSA a sales manager is easily able to determine areas to focus marketing efforts. If a particular FSA has consistently low volumes even though the assigned sales representative is consistently a high achiever the sales manager can direct the marketing department to focus their efforts on that particular area. One way of focusing marketing efforts on a specific geography is to rent a mail list from an Industry magazine and limit the list rental to only the targeted FSA.

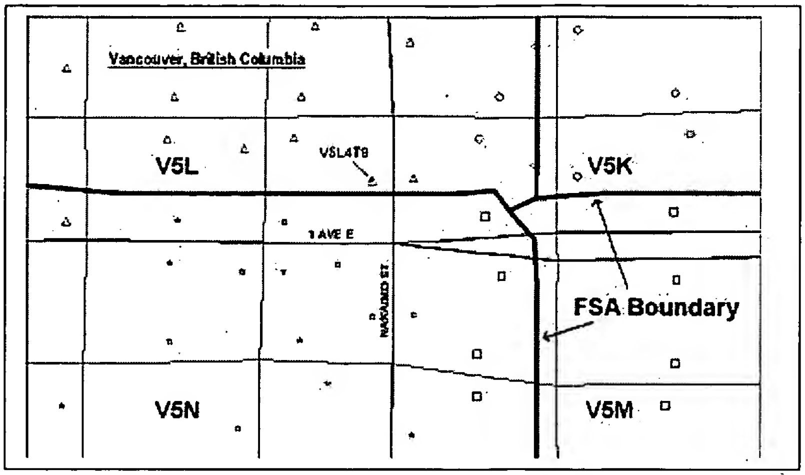


Diagram 2 - Traditional representation of FSA Boundaries digitally with geocoded postal code points shown

Direct marketers and service industries increasingly use Forward Sortation Area Boundaries to target their customers, prospects and retail site locations. By integrating their customers' FSA and transaction data, with FSA Boundaries, they can establish and analyze the best locations for their retail presence, provide service coverage details, reposition sales territories, or carry out targeted marketing campaigns. To do this they want the most accurate and up to date FSA Boundaries available.

Traditional FSA Boundaries

Traditional FSA Boundaries use streets to define the edges of the polygon boundaries. This is demonstrated in Diagram 1. Polygon boundaries defined only by street centerlines do not define exactly where the postal code points fall along the boundary line. This presents several problems especially if you intend to use FSA Boundaries as a means of geocoding your customers.

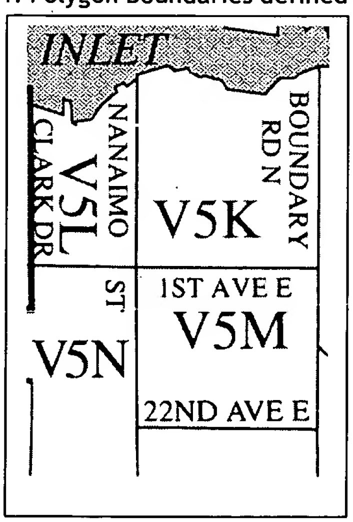


Diagram 1 - Traditional representation of FSA Boundaries

The largest challenge with traditional FSA Boundaries is the actual boundary lines themselves, as they do not indicate which side of the street belongs to which FSA. This is especially problematic when it comes to areas where several boundaries intersect. Think of a house on a corner where these boundaries intersect and already it becomes evident that by using traditional FSA Boundaries you could not say with certainty what the FSA for a particular house would be.

Straight FSA Boundaries along the street centerline such as those shown in Diagrams 1 and 2 do not cover postal codes on both sides of the street and they do not accurately account for houses or buildings that are near the intersection of two or more streets.

If accuracy is important to your application then traditional FSA Boundaries will not be sufficient.

The Next Generation of FSAs

To address the issues with traditional postal code boundary files, DMTI has developed the Street Intersection Data Model for FSAs.

Using CanMap Streetfiles as the base, DMTI Spatial systematically re-created the FSA Boundaries to account for where street addresses would actually fall on map. When encountering a street intersection, the

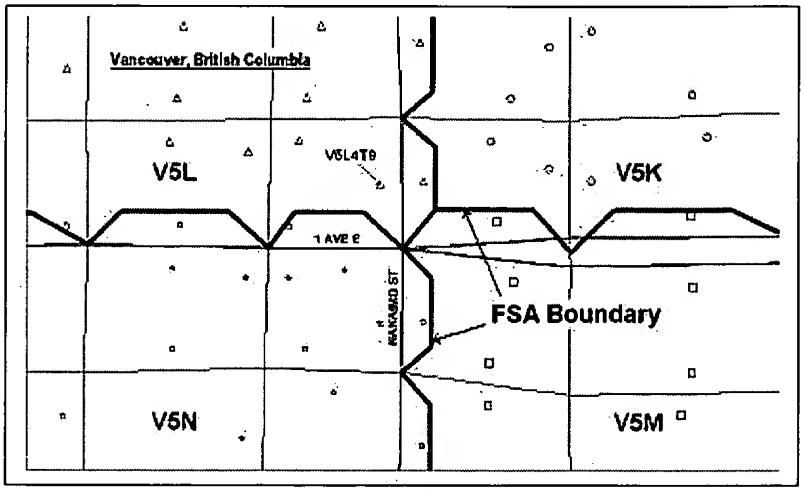


Diagram 3 - DMTI's Street Intersection Data Model for FSA Boundaries with geocoded postal code points shown.

boundaries intersect at approximately 45° degree angles to account for houses on the other intersecting block that would belong to a different FSA Boundary. Upon first glance these new boundaries may seem odd as they are scalloped along the entire boarder instead of the standard straight line boundaries that one may be used to.

You can see the difference that the new Street Intersection Data Model makes by looking at Diagrams 1, 2, and 3. Diagram 1 and 2 have the traditional street centreline boundaries. Diagram 2 is a digital representation of Diagram 1 with postal codes geocoded to the map and appearing as points represented by the triangles, circles, stars and squares. Diagram 3 has the same postal code points geocoded to the map as Diagram 2 however the FSA Boundaries are represented using the new Street Intersection Data Model first introduced by DMTI. In Diagram 2, which follows traditional FSA Boundaries, you can see that some of the postal codes represented by squares and circles are clearly contained within the wrong FSA Boundary. Whereas in Diagram 3 you can see that all of the appropriate postal codes are within the appropriate boundary. This makes for more accurate geocoding of your customers to the map, and ultimately more accurate analysis or use of the map.

The Street Intersection Data Model improved the geocoding hit rate for postal codes to 97.75%, an increase over the hit rate achieved by traditional FSA Boundaries which was 94.6%.

Interesting Facts About Postal Codes

- The most notorious FSA anomaly involves the Federal Government of Canada. Federal Government Buildings have a postal code beginning with K1A however, not all of the Federal Government offices actually fall within the K1A FSA Boundary on a map. Keep this in mind when geocoding as you may see stray postal codes in the wrong FSA which may appear to have geocoded incorrectly when in fact they are correct but it is an FSA anomaly.

- Rural postal codes can be distinguished from urban postal codes as the second character is "0" (zero).

- New Brunswick FSA Boundaries are in a state of change, the province has been undergoing the NB 9-1-1 Implementation Plan since 1998. This involves changing all rural addresses into civic addresses and changing all rural postal codes into Urban postal codes. In September 1998 E0C, E0H and E0J were converted, in March 1999 E0K and E0B were converted and in September 1999 E0G and E0E were converted from Rural to Urban (Civic addresses). The remainder of the conversion is scheduled for March 2000, at that time the entire province will be 9-1-1 ready and entirely covered by Urban FSAs. Canada Post will be expanding the rural to urban postal code conversion into other areas of the country over the next few years.

The DMTI Advantage

DMTI Spatial™ is Canada's premier publisher of precision build street map data (CanMap®)*, and innovative geocoding software (GeoPinpoint™)*. In addition, they publish a full range of positional accurate geo-spatial data products including: transportation & telecommunication data, census data and boundaries, postal geography, topographic maps, and marketing databases.

DMTI Spatial has experience in multiple industries including but not limited to: Real Estate, Telecommunications, Utilities, Government, Transportation, Banking, and Finance. Some of DMTI Spatial's clients include: Bell Canada, Enbridge Consumers Gas, Rogers Cablesystems, MapQuest, FedEx, Purolator, and the Department of National Defense.

Their data products are available in a variety of formats including MapInfo's .tab and .midmif and ESRI's .shp and .E00 formats. Other formats are available as custom orders.

DMTI Spatial's Service Department is able to provide clients with a comprehensive portfolio of services to complement and supplement your own in-house capabilities. These services include: GIS Consulting, Application development, Database Marketing, Data Conversion and Creation, Database Scrubbing, Geocoding Services, Technical Support, and Training Courses.

Responsive to the needs of their customers DMTI Spatial develops products with customer input in mind through a wish list program and error reporting via email directly to the product development team.

FSA Boundaries are updated twice per year in conjunction with the biannual release of the DMTI Spatial Enhanced Postal Code File*. With each update, new FSAs are added and retired FSAs are excluded. The Street Intersection Data Model for FSAs was first introduced to the FSA Boundary File by DMTI Spatial in December of 1999.

DMTI Spatial's Forward Sortation Area (FSA) Boundary File is available Nationwide in Unprojected latitude and longitude with the NAD83* Datum*.

DMTI Spatial has a Maintenance Program available for the FSA Boundary File which allows clients to subscribe in 1,2, or 3 year contracts. This convenient program allows clients to stay current with their data, which lets them focus on their work, knowing that they'll be using the most up-to-date GIS data. This is especially important with rural to urban postal code conversions taking place such as the one outlined for New Brunswick. As a program member, clients will automatically receive all updates within 30 days of the release date, allowing them to load the data immediately or whenever it fits into their project schedule.

In addition, DMTI Spatial will custom cut their data for the geography you require and will deliver the data via FTP site or CD-ROM.

Finding More Information

When purchasing data you want to make the most informed decision possible. For more sources of information check out the following:

- Industry trade shows and conferences
- Professional groups e.g., URISA (Urban and Regional Information Systems Association).
- User groups and special interest groups
- On-line user forums
- Industry news web sites: www.geoplace.com, www.giscafe.com, www.spatialnews.com, www.directionsmag.com
- Industry or system magazines (e.g., GeoWorld, GeoInfo Systems, Business Geographics, MapWorld, ArcNorth News)

Glossary of Terms and Products

Words denoted with a * are defined in the Glossary of Terms and Products.

Block-Face: refers to one side of a city street, normally between consecutive intersections with streets.

CanMap: from DMTI Spatial, is the world's number one choice for Canadian street map data. CanMap enables the user to carry out a range of sophisticated business geographic applications that require positional accuracy, detail, nationwide coverage, and

presentation quality cartographics.NAD83, Unprojected latitude, longitude. Maintenance subscription program available.

Centroid: the geographic center of any polygon.

Data Provider: a company that gather digital map data from a variety of public or private sources and adapts and enhances it for use within GIS application software for sales and marketing analysis of customers or prospects.

Datum: a mathematical model that provides a smooth approximation of the earth's surface. *See NAD*

Enhanced Postal Code File: from DMTI Spatial, is a point file representation of postal codes across Canada, with a geographic link to Statistics Canada's standard 1996 Census Boundaries.

Enumeration Area (EA): refers to the geographic area canvassed by one census representative. It is the smallest geographic area for which census data is reported. An Enumeration Area may contain approximately 125 to 440 dwellings depending on whether it is located in a rural or urban area respectively.

FSA: Forward Sortation Area. A polygon representing the first three characters of the Canadian Postal Code.

FSA LDU: Most commonly represented by a point, refers to the Canadian six digit postal code. FSA represents the first three digits of a postal code and LDU represents the last three digits of a postal code.

Postal Boundaries: *See FSA*

Geocode: to provide geographical coordinates for an address so that it can accurrately be placed on a map. *See GeoPinpoint.*

GeoPinpoint: DMTI Spatial software which attaches latitute and longitude geographical coordinates to your customer or prospect address data so that it can be accurately placed on a map.

GIS: Geographic Information System, a computer-based technology for retreiving, storing, and organizing data based on its location on a map.

NAD: North American Datum. Most current is NAD83 which was adopted by the Canadian Federal Government in 1990, and supersedes the North American Datum of 1927 (NAD27). *See Datum.*